

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Wassenegger, Michael
Riedel, Leonhard
Schiebel, Winfried
Sanger, Heinz

(ii) TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
POLYPEPTIDES HAVING THE ENZYMATIC ACTIVITY OF AN
RNA-DIRECTED RNA POLYMERASE (RdRP)

(iii) NUMBER OF SEQUENCES: 13

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: FISH & NEAVE
(B) STREET: 1251 Avenue of the Americas
(C) CITY: New York
(D) STATE: New York
(E) COUNTRY: USA
(F) ZIP: 10020

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/811,583
(B) FILING DATE: 05-MAR-1997
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Haley, James F.
(B) REGISTRATION NUMBER: 27,794
(C) REFERENCE/DOCKET NUMBER: MPG-1

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 212-596-9000
(B) TELEFAX: 212-596-9090

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3731 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Tomato

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 194..3535

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GAAATATTCT TTACTTACTT CACCAGGGAT TGACTCATCA CTCCCCCAA GTCTTTGTGT	60
GTTGTATAA TAAATTGGT TGTGCTTCAG TTTCAGTCAC TACTGCTGGG TAGTTTTAT	120
TTTGCATAAC TTCAGGGGGT ATTCCAGTTG GTGTTAGCAT TTGAAAGTCG AACTGCACTT	180
GGAATTGGC TAC ATG GGA AAG ACA ATT CAG GTT TTC GGA TTC CCT TAT	229
Met Gly Lys Thr Ile Gln Val Phe Gly Phe Pro Tyr	
1 5 10	
CTT CTC TCT GCG GAA GTG GTT AAG TCA TTC TTA GAG AAA TAT ACA GGA	277
Leu Leu Ser Ala Glu Val Val Lys Ser Phe Leu Glu Lys Tyr Thr Gly	
15 20 25	
TAT GGA ACT GTA TGT GCA TTG GAG GTT AAA CAG TCC AAA GGA GGA TCT	325
Tyr Gly Thr Val Cys Ala Leu Glu Val Lys Gln Ser Lys Gly Gly Ser	
30 35 40	
AGA GCA TTT GCC AAA GTT CAA TTT GCC GAC AAC ATA AGT GCT GAC AAA	373
Arg Ala Phe Ala Lys Val Gln Phe Ala Asp Asn Ile Ser Ala Asp Lys	
45 50 55 60	
ATC ATC ACT TTG GCT AAT AAC AGG CTG TAT TTT GGC TCT TCT TAT TTG	421
Ile Ile Thr Leu Ala Asn Asn Arg Leu Tyr Phe Gly Ser Ser Tyr Leu	
65 70 75	
AAG GCT TGG GAA ATG AAA ACT GAT ATT GTC CAA CTG CGG GCA TAT GTG	469
Lys Ala Trp Glu Met Lys Thr Asp Ile Val Gln Leu Arg Ala Tyr Val	
80 85 90	
GAT CAG ATG GAT GGC ATA ACT TTG AAT TTC GGA TGT CAG ATA TCA GAT	517
Asp Gln Met Asp Gly Ile Thr Leu Asn Phe Gly Cys Gln Ile Ser Asp	
95 100 105	
GAC AAG TTT GCA GTG TTG GGA AGT ACA GAA GTT TCA ATT CAA TTT GGC	565
Asp Lys Phe Ala Val Leu Gly Ser Thr Glu Val Ser Ile Gln Phe Gly	
110 115 120	
ATT GGA TTG AAG AAA TTT TTT TTC TTT TTA TCT AGT GGT TCA GCT GAC	613
Ile Gly Leu Lys Lys Phe Phe Phe Leu Ser Ser Gly Ser Ala Asp	
125 130 135 140	
TAT AAA CTT CAG CTT TCA TAT GAA AAT ATA TGG CAG GTT GTG CTC CAT	661
Tyr Lys Leu Gln Leu Ser Tyr Glu Asn Ile Trp Gln Val Val Leu His	
145 150 155	
CGT CCA TAT GGT CAA AAT GCT CAG TTT CTC CTC ATA CAG TTA TTT GGT	709
Arg Pro Tyr Gly Gln Asn Ala Gln Phe Leu Leu Ile Gln Leu Phe Gly	
160 165 170	
GCT CCT CGG ATC TAT AAG AGA CTT GAA AAC TCC TGT TAT AGC TTC TTT	757
Ala Pro Arg Ile Tyr Lys Arg Leu Glu Asn Ser Cys Tyr Ser Phe Phe	
175 180 185	
AAG GAA ACT CCT GAT GAT CAG TGG GTG AGG ACA ACA GAT TTC CCT CCA	805
Lys Glu Thr Pro Asp Asp Gln Trp Val Arg Thr Thr Asp Phe Pro Pro	
190 195 200	

TCT TGG ATA GGG CTA TCT TCT AGC TTA TGT TTG CAG TTC CGT AGG GGT Ser Trp Ile Gly Leu Ser Ser Ser Leu Cys Leu Gln Phe Arg Arg Gly 205 210 215 220	853
GTT CGT CTT CCA AAT TTC GAG GAA AGT TTT TTC CAC TAT GCA GAA CGT Val Arg Leu Pro Asn Phe Glu Glu Ser Phe Phe His Tyr Ala Glu Arg 225 230 235	901
GAA AAC AAT ATT ACT TTA CAG ACT GGT TTC ACC TTT TTC GTC TCT CAA Glu Asn Asn Ile Thr Leu Gln Thr Gly Phe Thr Phe Phe Val Ser Gln 240 245 250	949
AAA TCG GCT CTG GTT CCC AAT GTC CAG CCT CCG GAA GGA ATT TCA ATT Lys Ser Ala Leu Val Pro Asn Val Gln Pro Pro Glu Gly Ile Ser Ile 255 260 265	997
CCC TAC AAG ATT TTG TTC AAA ATT AGT TCT TTG GTA CAG CAT GGA TGC Pro Tyr Lys Ile Leu Phe Lys Ile Ser Ser Leu Val Gln His Gly Cys 270 275 280	1045
ATA CCT GGG CCA GCA TTA AAT GTC TAC TTT TTC CGA TTA GTT GAT CCT Ile Pro Gly Pro Ala Leu Asn Val Tyr Phe Phe Arg Leu Val Asp Pro 285 290 295 300	1093
CGA AGG AGA AAT GTG GCA TGC ATT GAG CAT GCC TTA GAG AAA CTG TAC Arg Arg Arg Asn Val Ala Cys Ile Glu His Ala Leu Glu Lys Leu Tyr 305 310 315	1141
TAT ATA AAG GAG TGC TGT TAT GAT CCC GTG AGG TGG CTC ACT GAG CAG Tyr Ile Lys Glu Cys Cys Tyr Asp Pro Val Arg Trp Leu Thr Glu Gln 320 325 330	1189
TAT GAT GGG TAT CTC AAG GGT AGA CAA CCT CCA AAA TCT CCG TCC ATC Tyr Asp Gly Tyr Leu Lys Gly Arg Gln Pro Pro Lys Ser Pro Ser Ile 335 340 345	1237
ACT TTA GAT GAT GGG TTG GTG TAT GTA AGA AGG GTC CTA GTA ACA CCA Thr Leu Asp Asp Gly Leu Val Tyr Val Arg Arg Val Leu Val Thr Pro 350 355 360	1285
TGC AAA GTT TAT TTT TGT GGT CCA GAG GTT AAT GTT TCC AAT CGG GTT Cys Lys Val Tyr Phe Cys Gly Pro Glu Val Asn Val Ser Asn Arg Val 365 370 375 380	1333
CTC CGC AAT TAT TCT GAA GAC ATA GAT AAC TTT CTT CGT GTT TCT TTT Leu Arg Asn Tyr Ser Glu Asp Ile Asp Asn Phe Leu Arg Val Ser Phe 385 390 395	1381
GTT GAT GAG GAG TGG GAG AAA CTG TAT TCT ACA GAC TTA TTA CCA AAA Val Asp Glu Glu Trp Glu Lys Leu Tyr Ser Thr Asp Leu Leu Pro Lys 400 405 410	1429
GCA AGT ACT GGA AGT GGT GTC AGG ACA AAC ATC TAT GAG AGG ATC TTA Ala Ser Thr Gly Ser Gly Val Arg Thr Asn Ile Tyr Glu Arg Ile Leu 415 420 425	1477
TCA ACT CTG CGG AAA GGC TTT GTA ATT GGT GAT AAA AAA TTT GAA TTT Ser Thr Leu Arg Lys Gly Phe Val Ile Gly Asp Lys Lys Phe Glu Phe 430 435 440	1525

CTT GCA TTT TCA TCG AGC CAG TTG CGG GAT AAT TCA GTG TGG ATG TTT Leu Ala Phe Ser Ser Gln Leu Arg Asp Asn Ser Val Trp Met Phe 445 450 455 460	1573
GCA TCA AGA CCT GGC CTT ACT GCA AAT GAT ATA AGA GCT TGG ATG GGT Ala Ser Arg Pro Gly Leu Thr Ala Asn Asp Ile Arg Ala Trp Met Gly 465 470 475	1621
GAT TTT TCG CAG ATC AAG AAT GTC GCA AAA TAT GCT GCC AGA CTT GGT Asp Phe Ser Gln Ile Lys Asn Val Ala Lys Tyr Ala Ala Arg Leu Gly 480 485 490	1669
CAA TCT TTT GGT TCC TCC AGA GAG ACT TTG AGT GTT CTT AGG CAT GAG Gln Ser Phe Gly Ser Ser Arg Glu Thr Leu Ser Val Leu Arg His Glu 495 500 505	1717
ATT GAA GTT ATT CCC GAT GTA AAG GTT CAT GGA ACC AGC TAT GTC TTT Ile Glu Val Ile Pro Asp Val Lys Val His Gly Thr Ser Tyr Val Phe 510 515 520	1765
TCT GAT GGA ATT GGT AAA ATA TCT GGT GAC TTT GCT CAT AGA GTT GCC Ser Asp Gly Ile Gly Lys Ile Ser Gly Asp Phe Ala His Arg Val Ala 525 530 535 540	1813
TCA AAA TGT GGC CTT CAA TAT ACC CCA TCT GCT TTC CAG ATT CGT TAT Ser Lys Cys Gly Leu Gln Tyr Thr Pro Ser Ala Phe Gln Ile Arg Tyr 545 550 555	1861
GGT GGA TAT AAA GGT GTT GTG GGT GTT GAT CCG GAT TCA TCA ATG AAG Gly Gly Tyr Lys Gly Val Val Gly Val Asp Pro Asp Ser Ser Met Lys 560 565 570	1909
TTG TCT TTG AGA AAG AGC ATG TCG AAA TAT GAA TCA GAC AAC ATA AAG Leu Ser Leu Arg Lys Ser Met Ser Lys Tyr Glu Ser Asp Asn Ile Lys 575 580 585	1957
TTA GAT GTC CTT GGA TGG AGC AAA TAT CAG CCT TGT TAT CTT AAT CGT Leu Asp Val Leu Gly Trp Ser Lys Tyr Gln Pro Cys Tyr Leu Asn Arg 590 595 600	2005
CAA CTG ATT ACG CTC TTG TCT ACA CTT GGA GTG AAA GAT GAA GTT CTC Gln Leu Ile Thr Leu Leu Ser Thr Leu Gly Val Lys Asp Glu Val Leu 605 610 615 620	2053
GAA CAG AAG CAA AAG GAA GCT GTA GAT CAG CTT GAT GCT ATC TTG CAT Glu Gln Lys Gln Lys Glu Ala Val Asp Gln Leu Asp Ala Ile Leu His 625 630 635	2101
GAT TCT TTG AAG GCA CAG GAG GCT TTG GAA TTG ATG TCT CCT GGA GAG Asp Ser Leu Lys Ala Gln Glu Ala Leu Glu Leu Met Ser Pro Gly Glu 640 645 650	2149
AAC ACT AAT ATT CTC AAG GCA ATG CTA AAC TGT GGT TAT AAG CCT GAT Asn Thr Asn Ile Leu Lys Ala Met Leu Asn Cys Gly Tyr Lys Pro Asp 655 660 665	2197
GCT GAG CCC TTT CTT TCA ATG ATG TTG CAA ACC TTC CGC GCA TCC AAG Ala Glu Pro Phe Leu Ser Met Met Leu Gln Thr Phe Arg Ala Ser Lys 670 675 680	2245

TTG CTC GAT TTG CGG ACT AGA TCA AGA ATA TTT ATT CCA AAT GGA AGA Leu Leu Asp Leu Arg Thr Arg Ser Arg Ile Phe Ile Pro Asn Gly Arg 685 690 695 700	2293
ACA ATG ATG GGA TGT TTG GAT GAA TCC AGA ACC TTG GAA TAT GGT CAG Thr Met Met Gly Cys Leu Asp Glu Ser Arg Thr Leu Glu Tyr Gly Gln 705 710 715	2341
G TG TTT GTT CAG TTT ACT GGT GCT GGA CAT GGA GAG TTT TCT GAC GAT Val Phe Val Gln Phe Thr Gly Ala Gly His Gly Glu Phe Ser Asp Asp 720 725 730	2389
TTA CAT CCA TTT AAT AAC AGC AGA TCC ACC AAC AGT AAT TTC ATT CTG Leu His Pro Phe Asn Asn Ser Arg Ser Thr Asn Ser Asn Phe Ile Leu 735 740 745	2437
AAG GGA AAT GTG GTT GCA AAA AAT CCA TGC TTG CAT CCT GGT GAT Lys Gly Asn Val Val Ala Lys Asn Pro Cys Leu His Pro Gly Asp 750 755 760	2485
ATT CGT GTT TTA AAG GCT GTA AAT GTT CGA GCG CTG CAC CAC ATG GTA Ile Arg Val Leu Lys Ala Val Asn Val Arg Ala Leu His His Met Val 765 770 775 780	2533
GAT TGT GTT GTA TTC CCT CAG AAA GGA AAA AGA CCT CAT CCG AAT GAA Asp Cys Val Val Phe Pro Gln Lys Gly Lys Arg Pro His Pro Asn Glu 785 790 795	2581
TGT TCT GGG AGT GAT TTG GAT GGG GAT ATC TAC TTT GTT TGC TGG GAT Cys Ser Gly Ser Asp Leu Asp Gly Asp Ile Tyr Phe Val Cys Trp Asp 800 805 810	2629
CAA GAC ATG ATC CCG CCA AGG CAA GTC CAG CCG ATG GAA TAT CCT CCA Gln Asp Met Ile Pro Pro Arg Gln Val Gln Pro Met Glu Tyr Pro Pro 815 820 825	2677
GCA CCC AGC ATA CAG TTG GAC CAT GAT GTC ACA ATT GAG GAA GTT GAA Ala Pro Ser Ile Gln Leu Asp His Asp Val Thr Ile Glu Glu Val Glu 830 835 840	2725
GAG TAC TTC ACC AAC TAT ATT GTG AAT GAC AGT TTG GGA ATC ATA GCA Glu Tyr Phe Thr Asn Tyr Ile Val Asn Asp Ser Leu Gly Ile Ile Ala 845 850 855 860	2773
AAT GCC CAT GTC GTA TTT GCA GAC AGA GAA CCT GAT ATG GCC ATG AGT Asn Ala His Val Val Phe Ala Asp Arg Glu Pro Asp Met Ala Met Ser 865 870 875	2821
GAT CCA TGC AAA AAA CTT GCT GAG CTC TTT TCA ATT GCA GTG GAC TTT Asp Pro Cys Lys Lys Leu Ala Glu Leu Phe Ser Ile Ala Val Asp Phe 880 885 890	2869
CCA AAG ACT GGT GTT CCC GCT GAA ATA CCA TCT CAG TTG CGC CCT AAA Pro Lys Thr Gly Val Pro Ala Glu Ile Pro Ser Gln Leu Arg Pro Lys 895 900 905	2917
GAA TAC CCA GAC TTC ATG GAT AAG CCG GAC AAG ACC AGC TAT ATC TCA Glu Tyr Pro Asp Phe Met Asp Lys Pro Asp Lys Thr Ser Tyr Ile Ser 910 915 920	2965

GAA AGA GTT ATT GGA AAG CTT TTC AGG AAA GTG AAG GAC AAA GCA CCT Glu Arg Val Ile Gly Lys Leu Phe Arg Lys Val Lys Asp Lys Ala Pro 925 930 935 940	3013
CAG GCT AGC TCT ATC GCG ACC TTC ACA AGA GAT GTT GCA AGG AGA TCA Gln Ala Ser Ser Ile Ala Thr Phe Thr Arg Asp Val Ala Arg Arg Ser 945 950 955	3061
TAT GAT GCT GAT ATG GAA GTT GAT GGA TTT GAA GAT TAC ATT GAC GAA Tyr Asp Ala Asp Met Glu Val Asp Gly Phe Glu Asp Tyr Ile Asp Glu 960 965 970	3109
GCT TTT GAC TAC AAA ACT GAA TAT GAC AAC AAG CTG GGT AAT TTA ATG Ala Phe Asp Tyr Lys Thr Glu Tyr Asp Asn Lys Leu Gly Asn Leu Met 975 980 985	3157
GAC TAC TAT GGC ATA AAA ACA GAG GCT GAA ATA CTT AGT GGT GGC ATT Asp Tyr Tyr Gly Ile Lys Thr Glu Ala Glu Ile Leu Ser Gly Gly Ile 990 995 1000	3205
ATG AAG GCA TCA AAA ACT TTT GAC CGC AGA AAA GAT GCT GAG GCC ATT Met Lys Ala Ser Lys Thr Phe Asp Arg Arg Lys Asp Ala Glu Ala Ile 1005 1010 1015 1020	3253
AGT GTT GCT GTG AGG GCC TTG AGG AAG GAG GCA AGA GCC TGG TTC AAG Ser Val Ala Val Arg Ala Leu Arg Lys Glu Ala Arg Ala Trp Phe Lys 1025 1030 1035	3301
AGG CGT AAT GAT ATA GAT GAC ATG TTA CCA AAG GCT TCG GCT TGG TAC Arg Arg Asn Asp Ile Asp Asp Met Leu Pro Lys Ala Ser Ala Trp Tyr 1040 1045 1050	3349
CAC GTT ACA TAT CAT CCT ACA TAT TGG GGT TGC TAC AAT CAG GGG TTG His Val Thr Tyr His Pro Thr Tyr Trp Gly Cys Tyr Asn Gln Gly Leu 1055 1060 1065	3397
AAA AGA GCT CAT TTC ATT AGC TTT CCC TGG TGT GTT TAT GAC CAG CTA Lys Arg Ala His Phe Ile Ser Phe Pro Trp Cys Val Tyr Asp Gln Leu 1070 1075 1080	3445
ATC CAG ATT AAG AAG GAC AAA GCA CGT AAC AGG CCA GTT CTC AAC TTG Ile Gln Ile Lys Lys Asp Lys Ala Arg Asn Arg Pro Val Leu Asn Leu 1085 1090 1095 1100	3493
TCA TCT CTC AGG GCT CAA CTG AGT CAC AGA TTA GTG TTG AAA Ser Ser Leu Arg Ala Gln Leu Ser His Arg Leu Val Leu Lys 1105 1110	3535
TGAGATTCCA GTCGAGCGTT AAGCTGATAT ATATATAATG TAATAGGGTG TGATCATAAG	3595
AAAAACTGTTA TGCATTGTTG ACTACCTTT GTCTTTAAAAA CTGCATGAAG CTGCAACATA	3655
TATGCAGTAC TCTAAGAAC AGATGTACAG CTAAGTACTA ATATGTATGT GATTTGAGTT	3715
TCATCTTCT TCTAAA	3731

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1114 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Gly	Lys	Thr	Ile	Gln	Val	Phe	Gly	Phe	Pro	Tyr	Leu	Leu	Ser	Ala
1				5					10					15	
Glu	Val	Val	Lys	Ser	Phe	Leu	Glu	Lys	Tyr	Thr	Gly	Tyr	Gly	Thr	Val
			20					25					30		
Cys	Ala	Leu	Glu	Val	Lys	Gln	Ser	Lys	Gly	Gly	Ser	Arg	Ala	Phe	Ala
			35				40					45			
Lys	Val	Gln	Phe	Ala	Asp	Asn	Ile	Ser	Ala	Asp	Lys	Ile	Ile	Thr	Leu
			50			55					60				
Ala	Asn	Asn	Arg	Leu	Tyr	Phe	Gly	Ser	Ser	Tyr	Leu	Lys	Ala	Trp	Glu
	65			70				75					80		
Met	Lys	Thr	Asp	Ile	Val	Gln	Leu	Arg	Ala	Tyr	Val	Asp	Gln	Met	Asp
				85				90					95		
Gly	Ile	Thr	Leu	Asn	Phe	Gly	Cys	Gln	Ile	Ser	Asp	Asp	Lys	Phe	Ala
			100				105						110		
Val	Leu	Gly	Ser	Thr	Glu	Val	Ser	Ile	Gln	Phe	Gly	Ile	Gly	Leu	Lys
	115				120							125			
Lys	Phe	Phe	Phe	Leu	Ser	Ser	Gly	Ser	Ala	Asp	Tyr	Lys	Leu	Gln	
	130			135				140							
Leu	Ser	Tyr	Glu	Asn	Ile	Trp	Gln	Val	Val	Leu	His	Arg	Pro	Tyr	Gly
	145				150				155			160			
Gln	Asn	Ala	Gln	Phe	Leu	Leu	Ile	Gln	Leu	Phe	Gly	Ala	Pro	Arg	Ile
			165				170					175			
Tyr	Lys	Arg	Leu	Glu	Asn	Ser	Cys	Tyr	Ser	Phe	Phe	Lys	Glu	Thr	Pro
			180				185					190			
Asp	Asp	Gln	Trp	Val	Arg	Thr	Thr	Asp	Phe	Pro	Pro	Ser	Trp	Ile	Gly
			195			200						205			
Leu	Ser	Ser	Ser	Leu	Cys	Leu	Gln	Phe	Arg	Arg	Gly	Val	Arg	Leu	Pro
	210				215				220						
Asn	Phe	Glu	Glu	Ser	Phe	Phe	His	Tyr	Ala	Glu	Arg	Glu	Asn	Asn	Ile
	225				230				235				240		
Thr	Leu	Gln	Thr	Gly	Phe	Thr	Phe	Phe	Val	Ser	Gln	Lys	Ser	Ala	Leu
			245				250					255			
Val	Pro	Asn	Val	Gln	Pro	Pro	Glu	Gly	Ile	Ser	Ile	Pro	Tyr	Lys	Ile
			260				265					270			
Leu	Phe	Lys	Ile	Ser	Ser	Leu	Val	Gln	His	Gly	Cys	Ile	Pro	Gly	Pro
			275				280					285			

Ala Leu Asn Val Tyr Phe Phe Arg Leu Val Asp Pro Arg Arg Arg Asn
 290 295 300
 Val Ala Cys Ile Glu His Ala Leu Glu Lys Leu Tyr Tyr Ile Lys Glu
 305 310 315 320
 Cys Cys Tyr Asp Pro Val Arg Trp Leu Thr Glu Gln Tyr Asp Gly Tyr
 325 330 335
 Leu Lys Gly Arg Gln Pro Pro Lys Ser Pro Ser Ile Thr Leu Asp Asp
 340 345 350
 Gly Leu Val Tyr Val Arg Arg Val Leu Val Thr Pro Cys Lys Val Tyr
 355 360 365
 Phe Cys Gly Pro Glu Val Asn Val Ser Asn Arg Val Leu Arg Asn Tyr
 370 375 380
 Ser Glu Asp Ile Asp Asn Phe Leu Arg Val Ser Phe Val Asp Glu Glu
 385 390 395 400
 Trp Glu Lys Leu Tyr Ser Thr Asp Leu Leu Pro Lys Ala Ser Thr Gly
 405 410 415
 Ser Gly Val Arg Thr Asn Ile Tyr Glu Arg Ile Leu Ser Thr Leu Arg
 420 425 430
 Lys Gly Phe Val Ile Gly Asp Lys Lys Phe Glu Phe Leu Ala Phe Ser
 435 440 445
 Ser Ser Gln Leu Arg Asp Asn Ser Val Trp Met Phe Ala Ser Arg Pro
 450 455 460
 Gly Leu Thr Ala Asn Asp Ile Arg Ala Trp Met Gly Asp Phe Ser Gln
 465 470 475 480
 Ile Lys Asn Val Ala Lys Tyr Ala Ala Arg Leu Gly Gln Ser Phe Gly
 485 490 495
 Ser Ser Arg Glu Thr Leu Ser Val Leu Arg His Glu Ile Glu Val Ile
 500 505 510
 Pro Asp Val Lys Val His Gly Thr Ser Tyr Val Phe Ser Asp Gly Ile
 515 520 525
 Gly Lys Ile Ser Gly Asp Phe Ala His Arg Val Ala Ser Lys Cys Gly
 530 535 540
 Leu Gln Tyr Thr Pro Ser Ala Phe Gln Ile Arg Tyr Gly Gly Tyr Lys
 545 550 555 560
 Gly Val Val Gly Val Asp Pro Asp Ser Ser Met Lys Leu Ser Leu Arg
 565 570 575
 Lys Ser Met Ser Lys Tyr Glu Ser Asp Asn Ile Lys Leu Asp Val Leu
 580 585 590
 Gly Trp Ser Lys Tyr Gln Pro Cys Tyr Leu Asn Arg Gln Leu Ile Thr
 595 600 605

Leu Leu Ser Thr Leu Gly Val Lys Asp Glu Val Leu Glu Gln Lys Gln
 610 615 620

Lys Glu Ala Val Asp Gln Leu Asp Ala Ile Leu His Asp Ser Leu Lys
 625 630 635 640

Ala Gln Glu Ala Leu Glu Leu Met Ser Pro Gly Glu Asn Thr Asn Ile
 645 650 655

Leu Lys Ala Met Leu Asn Cys Gly Tyr Lys Pro Asp Ala Glu Pro Phe
 660 665 670

Leu Ser Met Met Leu Gln Thr Phe Arg Ala Ser Lys Leu Leu Asp Leu
 675 680 685

Arg Thr Arg Ser Arg Ile Phe Ile Pro Asn Gly Arg Thr Met Met Gly
 690 695 700

Cys Leu Asp Glu Ser Arg Thr Leu Glu Tyr Gly Gln Val Phe Val Gln
 705 710 715 720

Phe Thr Gly Ala Gly His Gly Glu Phe Ser Asp Asp Leu His Pro Phe
 725 730 735

Asn Asn Ser Arg Ser Thr Asn Ser Asn Phe Ile Leu Lys Gly Asn Val
 740 745 750

Val Val Ala Lys Asn Pro Cys Leu His Pro Gly Asp Ile Arg Val Leu
 755 760 765

Lys Ala Val Asn Val Arg Ala Leu His His Met Val Asp Cys Val Val
 770 775 780

Phe Pro Gln Lys Gly Lys Arg Pro His Pro Asn Glu Cys Ser Gly Ser
 785 790 795 800

Asp Leu Asp Gly Asp Ile Tyr Phe Val Cys Trp Asp Gln Asp Met Ile
 805 810 815

Pro Pro Arg Gln Val Gln Pro Met Glu Tyr Pro Pro Ala Pro Ser Ile
 820 825 830

Gln Leu Asp His Asp Val Thr Ile Glu Glu Val Glu Tyr Phe Thr
 835 840 845

Asn Tyr Ile Val Asn Asp Ser Leu Gly Ile Ile Ala Asn Ala His Val
 850 855 860

Val Phe Ala Asp Arg Glu Pro Asp Met Ala Met Ser Asp Pro Cys Lys
 865 870 875 880

Lys Leu Ala Glu Leu Phe Ser Ile Ala Val Asp Phe Pro Lys Thr Gly
 885 890 895

Val Pro Ala Glu Ile Pro Ser Gln Leu Arg Pro Lys Glu Tyr Pro Asp
 900 905 910

Phe Met Asp Lys Pro Asp Lys Thr Ser Tyr Ile Ser Glu Arg Val Ile
 915 920 925

Gly Lys Leu Phe Arg Lys Val Lys Asp Lys Ala Pro Gln Ala Ser Ser
 930 935 940

Ile Ala Thr Phe Thr Arg Asp Val Ala Arg Arg Ser Tyr Asp Ala Asp
 945 950 955 960

Met Glu Val Asp Gly Phe Glu Asp Tyr Ile Asp Glu Ala Phe Asp Tyr
 965 970 975

Lys Thr Glu Tyr Asp Asn Lys Leu Gly Asn Leu Met Asp Tyr Tyr Gly
 980 985 990

Ile Lys Thr Glu Ala Glu Ile Leu Ser Gly Gly Ile Met Lys Ala Ser
 995 1000 1005

Lys Thr Phe Asp Arg Arg Lys Asp Ala Glu Ala Ile Ser Val Ala Val
 1010 1015 1020

Arg Ala Leu Arg Lys Glu Ala Arg Ala Trp Phe Lys Arg Arg Asn Asp
 1025 1030 1035 1040

Ile Asp Asp Met Leu Pro Lys Ala Ser Ala Trp Tyr His Val Thr Tyr
 1045 1050 1055

His Pro Thr Tyr Trp Gly Cys Tyr Asn Gln Gly Leu Lys Arg Ala His
 1060 1065 1070

Phe Ile Ser Phe Pro Trp Cys Val Tyr Asp Gln Leu Ile Gln Ile Lys
 1075 1080 1085

Lys Asp Lys Ala Arg Asn Arg Pro Val Leu Asn Leu Ser Ser Leu Arg
 1090 1095 1100

Ala Gln Leu Ser His Arg Leu Val Leu Lys
 1105 1110

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 218 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Arg Thr Met Met Gly Cys Leu Asp Glu Ser Arg Thr Leu Glu Tyr Gly
 1 5 10 15

Gln Val Phe Val Gln Phe Thr Gly Ala Gly His Gly Glu Phe Ser Asp
 20 25 30

Asp Leu His Pro Phe Asn Asn Ser Arg Ser Thr Asn Ser Asn Phe Ile
 35 40 45

Leu Lys Gly Asn Val Val Val Ala Lys Asn Pro Cys Leu His Pro Gly
 50 55 60

Asp Ile Arg Val Leu Lys Ala Val Asn Val Arg Ala Leu His His Met
 65 70 75 80

Val Asp Cys Val Val Phe Pro Gln Lys Gly Lys Arg Pro His Pro Asn
 85 90 95

Glu Cys Ser Gly Ser Asp Leu Asp Gly Asp Ile Tyr Phe Val Cys Trp
 100 105 110

Asp Gln Asp Met Ile Pro Pro Arg Gln Val Gln Pro Met Glu Tyr Pro
 115 120 125

Pro Ala Pro Ser Ile Gln Leu Asp His Asp Val Thr Ile Glu Glu Val
 130 135 140

Glu Glu Tyr Phe Thr Asn Tyr Ile Val Asn Asp Ser Leu Gly Ile Ile
 145 150 155 160

Ala Asn Ala His Val Val Phe Ala Asp Arg Glu Pro Asp Met Ala Met
 165 170 175

Ser Asp Pro Cys Lys Lys Leu Ala Glu Leu Phe Ser Ile Ala Val Asp
 180 185 190

Phe Pro Lys Thr Gly Val Pro Ala Glu Ile Pro Ser Gln Leu Arg Pro
 195 200 205

Lys Glu Tyr Pro Asp Phe Met Asp Lys Pro
 210 215

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CATAACGAAT CTGGAAAGCA GATGG

25

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GATGAATCCG GATCAACACC CACAC

25

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGGTGCTGGA GGATATTCCA TCGGC

25

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTTCACCAGG GATCCACTCA TCACCTCCCT CAAG

34

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCATAACTTC AGGGGGGATC CAGTTGGTGT TAGC

34

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCAGCTTCAT GCAGATCTAA AGACAAAAAGG TAGTC

35

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 14 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Ser Asn Arg Val Leu Arg Asn Tyr Ser Glu Asp Ile Asp Asn
1 5 10

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 14 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Ala Ser Lys Thr Phe Asp Arg Arg Lys Asp Ala Glu Ala Ile
1 5 10

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Glu Gln Tyr Asp Gly Tyr Leu Lys Gly Arg Gln Pro Pro Lys Ser Pro
1 5 10 15
Ser

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Val Phe Pro Gln Lys Gly Lys Arg Pro His Asn Glu Cys
1 5 10